



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/664,044
Source: IFwo
Date Processed by STIC: 2/17/04

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221

Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two:
2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office,
Box Sequence, Room 4B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: <u>10/664,044</u>
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 <input checked="" type="checkbox"/> Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 <input type="checkbox"/> Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.	
3 <input checked="" type="checkbox"/> Misaligned Amino Numbering	The numbering under each 5 th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.	
4 <input type="checkbox"/> Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.	
5 <input type="checkbox"/> Variable Length	Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 <input type="checkbox"/> PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.	
7 <input type="checkbox"/> Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION: SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.	
8 <input type="checkbox"/> Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000	
9 <input type="checkbox"/> Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
10 <input type="checkbox"/> Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence	
11 <input type="checkbox"/> Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 <input type="checkbox"/> PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 <input type="checkbox"/> Misuse of n/Xaa	"n" can only represent a single <u>nucleotide</u> ; "Xaa" can only represent a single <u>amino acid</u>	



IFWO

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/664,044

DATE: 02/17/2004

TIME: 12:17:46

Input Set : A:\PTO.YF.txt

Output Set: N:\CRF4\02172004\J664044.raw

2 <110> APPLICANT: Japan Atomic Energy Research Institute
W--> 3 <120> TITLE OF INVENTION: A method for efficiently determining a DNA strand
W--> 4 break
W--> 5 <130> FILE REFERENCE: 030217
C--> 6 <140> CURRENT APPLICATION NUMBER: US/10/664,044
C--> 6 <141> CURRENT FILING DATE: 2003-09-17
W--> 6 <160> NUMBER OF SEQ ID: 4
E--> 8 <200> 1

*Please consult
Sequence Rules.*

*Does Not Comply
Corrected Diskette Needed*

ERRORED SEQUENCES

9 <211> LENGTH: 284 *855*
10 <212> TYPE: *PRT* *DNA*
11 <213> ORGANISM: Deinococcus radiodurans, strain KD8301
W--> 12 <220> FEATURE:
13 <223> OTHER INFORMATION: *Amino acid sequence* of DNA repair promoting protein, *sequence*.
14 PprA, of Deinococcus radiodurans, strain KD8301.

E--> 16 <400> SEQUENCE: 1
E--> 17 atg gca agg gct aaa gca aaa gac caa acg gac ggc atc tac gcc gcc
E--> 18 48
19 Met Ala Arg Ala Lys Ala Lys Asp Gln Thr Asp Gly Ile Tyr Ala Ala
E--> 20 1 5 10 15
E--> 21 ttc gac acc ttg atg agc acg gcg ggc gtg gac agc cag atc gcc gcc
E--> 22 96
23 Phe Asp Thr Leu Met Ser Thr Ala Gly Val Asp Ser Gln Ile Ala Ala
E--> 24 20 25 30
E--> 25 ctc gcc gcg agt gag gcc gac gcg ggc acg ctg gac gcg gcg ctc acg
E--> 26 144
27 Leu Ala Ala Ser Glu Ala Asp Ala Gly Thr Leu Asp Ala Ala Leu Thr
E--> 28 35 40 45
E--> 29 cag tcc ttg caa gaa gcg cag ggg cgc tgg ggg ctg ggg ctg cac cac
E--> 30 192
31 Gln Ser Leu Gln Glu Ala Gln Gly Arg Trp Gly Leu Gly Leu His His
E--> 32 50 55 60
E--> 33 ctg cgc cat gag gcg cgg ctg acc gac gac ggc gac atc gaa att ctg
E--> 34 240
35 Leu Arg His Glu Ala Arg Leu Thr Asp Asp Gly Asp Ile Glu Ile Leu
E--> 36 65 70 75
E--> 37 80
E--> 38 acc gat ggc cgc ccc agc gcc cgc gtg agc gag ggc ttc gga gca ctc
E--> 39 288

*This sequence is not
an amino acid
sequence.*

*see item 1 on
48 Error Summary
sheet.*

*do not use TAB codes
between
amino acid
numbers
(see item 3 on
Error Summary
sheet)*

*A coding
sequence is
not a PRT
sequence; it is
a nucleotide
sequence.*

RAW SEQUENCE LISTING

DATE: 02/17/2004

PATENT APPLICATION: US/10/664,044

TIME: 12:17:46

Input Set : A:\PTO.YF.txt

Output Set: N:\CRF4\02172004\J664044.raw

```

      40 Thr Asp Gly Arg Pro Ser Ala Arg Val Ser Glu Gly Phe Gly Ala Leu
E--> 41      85      90      95
E--> 42 gcg cag gcc tac gcg ccc atg cag gcg ctc gac gaa cgc ggc ctg agc
E--> 43 336
      44 Ala Gln Ala Tyr Ala Pro Met Gln Ala Leu Asp Glu Arg Gly Leu Ser
E--> 45      100      105      110
E--> 46 cag tgg gcg gcg ctc ggc gag ggc tac cgc gct ccc ggc gac ttg ccg
E--> 47 384
      48 Gln Trp Ala Ala Leu Gly Glu Gly Tyr Arg Ala Pro Gly Asp Leu Pro
E--> 49      115      120      125
E--> 50 ttg gcg cag ctc aag gtg ctg atc gag cac gcc cgc gac ttc gaa acc
E--> 51 432
      52 Leu Ala Gln Leu Lys Val Leu Ile Glu His Ala Arg Asp Phe Glu Thr
E--> 53      130      135      140
E--> 54 gac tgg tcg gcg ggg cgc ggc gaa acc ttt cag cgc gtg tgg cgc aag
E--> 55 480
      56 Asp Trp Ser Ala Gly Arg Gly Glu Thr Phe Gln Arg Val Trp Arg Lys
E--> 57 145      150      155
E--> 58 160
E--> 59 ggc gac acc ctg ttt gtc gag gtg gcc cgg ccc gcg tcc gcc gag gcc
E--> 60 528
      61 Gly Asp Thr Leu Phe Val Glu Val Ala Arg Pro Ala Ser Ala Glu Ala
E--> 62      165      170      175
E--> 63 gcg ctc tcc gac gct gcc tgg gac gtg atc gcc agc atc aag gac cgc
E--> 64 576
      65 Ala Leu Ser Asp Ala Ala Trp Asp Val Ile Ala Ser Ile Lys Asp Arg
E--> 66      180      185      190
E--> 67 gcc ttc cag cgt gag ctg atg cgc cgc agc gag aag gac ggg atg ctc
E--> 68 624
      69 Ala Phe Gln Arg Glu Leu Met Arg Arg Ser Glu Lys Asp Gly Met Leu
E--> 70      195      200      205
E--> 71 ggc gcc ctg ctc ggg gct cgc cac gcc ggg gcc aag gcc aac ctc gcc
E--> 72 672
      73 Gly Ala Leu Leu Gly Ala Arg His Ala Gly Ala Lys Ala Asn Leu Ala
E--> 74      210      215      220
E--> 75 cag ctg ccc gaa gcg cac ttc acc gtg cag gcg ttc gtg cag acc ctc
E--> 76 720
      77 Gln Leu Pro Glu Ala His Phe Thr Val Gln Ala Phe Val Gln Thr Leu
E--> 78 225      230      235
E--> 79 240
E--> 80 agc gga gcc gcc gcc cgc aac gcc gag gag tac cgc gcg gcc ctg aaa
E--> 81 768
      82 Ser Gly Ala Ala Ala Arg Asn Ala Glu Glu Tyr Arg Ala Ala Leu Lys
E--> 83      245      250      255
E--> 84 acc gcc gcc gct gcg ctg gag gaa tac cag ggc gtg acc acc cgc caa
E--> 85 816
      86 Thr Ala Ala Ala Ala Leu Glu Glu Tyr Gln Gly Val Thr Thr Arg Gln
E--> 87      260      265      270
E--> 88 ctg tcc gaa gtg ctg cgg cac ggc ctg cgc gag agc tga

```

*same
error
as
page 1*

RAW SEQUENCE LISTING

DATE: 02/17/2004

PATENT APPLICATION: US/10/664,044

TIME: 12:17:46

Input Set : A:\PTO.YF.txt

Output Set: N:\CRF4\02172004\J664044.raw

E--> 89 855

E--> 90 Leu Ser Glu Val Leu Arg His Gly Leu Arg Glu Ser Sto

E--> 91 275 280 285

E--> 93 <200> 2 <2107>

94 <211> LENGTH: 855

95 <212> TYPE: DNA

96 <213> ORGANISM: Deinococcus radiodurans, strain KD8301

W--> 97 <220> FEATURE:

98 <223> OTHER INFORMATION: Nucleotide sequence of DNA repair promoting protein,

99 pprA, of Deinococcus radiodurans, strain KD8301.

E--> 101 <400> SEQUENCE: 2

102 atggcaaggg ctaaagcaaa agaccaaagc gacggcatct acgccgcctt cgacaccttg 60

103 atgagcacgg cgggcgtgga cagccagatc gccgcctctg ccgcgagtga ggccgacgcg 120

104 ggcacgctgg acgcggcgct cagcagtcct ttgcaagaag cgcagggggc ctgggggctg 180

105 gggctgcacc acctgcgcca tgaggcgcgg ctgaccgacg acggcgacat cgaaattctg 240

106 accgatggcc gccccagcgc ccgcgtgagc gagggttcg gagcactcgc gcaggcctac 300

107 gcgcccattg aggcgctcga cgaacgcggc ctgagccagt gggcgggcgt cggcgagggc 360

108 taccgcgctc ccggcgactt gccgttggcg cagctcaagg tgctgatcga gcacgccgcg 420

109 gacttcgaaa ccgactggtc ggcgggggcg ggcgaaacct ttcagcgctg gtggcgcaag 480

110 ggcyacaccc tgtttgtcga ggtggcccg ccgcgctccg ccgaggccgc gctctccgac 540

111 gctgcctggg acgtgatcgc cagcatcaag gaccgcgctt tccagcgtga gctgatgcgc 600

112 cgcagcgaga aggacgggat gctcggcgcc ctgctcgggg ctgccacgc cggggccaag 660

113 gccaacctcg ccagctgcc cgaagcgac ttcaccgtgc aggcgttcgt gcagaccctc 720

114 agcggagccg ccgccgcaa cgccgaggag taccgcgcgg cctgaaaac cgccgccgct 780

115 gcgctggagg aataccaggg cgtgaccacc cgccaactgt ccgaagtgt gcggcacggc 840

E--> 116 ctgcgcgaga gctga

117 855

E--> 119 <200> 3 <2107>

120 <211> LENGTH: 35

121 <212> TYPE: DNA

122 <213> ORGANISM: Artificial sequence

W--> 123 <220> FEATURE:

124 <223> OTHER INFORMATION: Sense primer for amplifying pprA gene.

E--> 126 <400> SEQUENCE: 3

E--> 127 gggcataata aaggccatat ggcaagggct aaagc

128 35

E--> 130 <200> 4 <2107>

131 <211> LENGTH: 32

132 <212> TYPE: DNA

133 <213> ORGANISM: Artificial sequence

W--> 134 <220> FEATURE:

135 <223> OTHER INFORMATION: Antisense primer for amplifying pprA gene.

E--> 137 <400> SEQUENCE: 4

E--> 138 ttttggatcc tcagctctcg cgcaggccgt gc

139 32

855 (see item 1
on Error
Summary
Sheet)

35 (see item 1)

32 (see item 1)

Please consult sample Sequence Listing
(attached in back)

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/664,044

DATE: 02/17/2004
TIME: 12:17:47

Input Set : A:\PTO.YF.txt
Output Set: N:\CRF4\02172004\J664044.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:2; Line(s) 116

VERIFICATION SUMMARY

DATE: 02/17/2004

PATENT APPLICATION: US/10/664,044

TIME: 12:17:47

Input Set : A:\PTO.YF.txt

Output Set: N:\CRF4\02172004\J664044.raw

L:3 M:283 W: Missing Blank Line separator, <120> field identifier
L:5 M:283 W: Missing Blank Line separator, <130> field identifier
L:6 M:270 C: Current Application Number differs, Replaced Current Application No
L:6 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:6 M:283 W: Missing Blank Line separator, <160> field identifier
L:8 M:250 E: Invalid Numeric Identifier, INVALID IDENTIFIER
L:12 M:283 W: Missing Blank Line separator, <220> field identifier
L:16 M:282 E: Numeric Field Identifier Missing, <210> is required.
L:16 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:0 differs:1
L:17 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:18 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:1
M:332 Repeated in SeqNo=33
L:21 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:25 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:29 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:33 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:38 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:42 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:46 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:50 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:54 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:59 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:63 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:67 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:71 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:75 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:80 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:84 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:16
L:88 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:13
L:90 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
L:91 M:252 E: No. of Seq. differs, <211> LENGTH:Input:284 Found:570 SEQ:0
L:93 M:250 E: Invalid Numeric Identifier, INVALID IDENTIFIER
L:97 M:283 W: Missing Blank Line separator, <220> field identifier
L:101 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:0 differs:2
L:116 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:855 SEQ:2
L:119 M:250 E: Invalid Numeric Identifier, INVALID IDENTIFIER
L:123 M:283 W: Missing Blank Line separator, <220> field identifier
L:126 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:0 differs:3
M:254 Repeated in SeqNo=33
L:130 M:250 E: Invalid Numeric Identifier, INVALID IDENTIFIER
L:134 M:283 W: Missing Blank Line separator, <220> field identifier
L:137 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:0 differs:4

<110> Smith, John; Smithgene Inc.

<120> Example of a Sequence Listing

<130> 01-00001

<140> PCT/EP98/00001
<141> 1998-12-31

<150> US 08/999,999
<151> 1997-10-15

<160> <

<170> PatentIn version 2.0

<210> 1
<211> 389
<212> DNA
<213> Paramecium sp.

<220>
<221> CDS
<222> (279)...(389)

<300>
<301> Doc. Richard
<302> Isolation and Characterization of a Gene Encoding a
Protease from Paramecium sp.
<303> Journal of Genes
<304> 1
<305> 4
<306> 1-7
<307> 1988-06-31
<308> 123456
<309> 1980-06-31

<400> 1
agctgtagtc attcctgtgt cctctctctc ctgggcttct cccccctgcta atcagatctc 60
agggagagtg tcttgacctt cctctgcctt tgcagcttca caggcaggca ggcaggcagc 120
tgatgtggca attgctggca gtgccacagg ctttctcagcc aggccttaggg tgggttccgc 180
cgcgggcgcg cgccctctct cgcgctctct tcgcgctctt ctctcgtctt cctctcgtct 240

Consult this

Appendix 3, page 2

ggacctgatt aggtgagcag gaggagggggg cagcttagc 290
Met Val Ser Met Phe Ser

ctg tct ttc aaa tgg cct gga tct tgt ttg tct gtt tgt ttg ttc caa 300
Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu Phe Val Cys Leu Phe Cln

tgt ccc aaa gtc ttc ccc tgt cac tca tca ctg cag ccg aat cct 389
Cys Pro Lys Val Leu Pro Cys His Ser Ser Leu Gln Pro Asn Leu

<210> 2
<211> 37
<212> PRT
<213> Paramecium sp.

<<00> 2
Met Val Ser Met Phe Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu
1 5 10 15

Phe Val Cys Leu Phe Cln Cys Pro Lys Val Leu Pro Cys His Ser Ser
20 25 30

Leu Cln Pro Asn Leu
35

<210> 3
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Designed peptide based on size and polarity to act as a linker between the alpha and beta chains of Protein XYZ.

<400> 3
Met Val Asn Leu Glu Pro Met His Thr Glu Ile
1 5 10

<210> 4
<400> 4
000

[Annex VIII follows]

identifiers and their accompanying information as shown in the following table. The numeric identifier shall be used only in the "Sequence Listing." The order and presentation of the items of information in the "Sequence Listing" shall conform to the arrangement given below. Each item of information shall begin on a new line and shall begin with the numeric identifier enclosed in angle brackets as shown. The submission of those items of information designated with an "M" is mandatory. The submission of those items of information designated with an "O" is optional. Numeric identifiers <110> through <170> shall only be set forth at the beginning of the "Sequence Listing." The following table illustrates the numeric identifiers.

Numeric Identifier	Definition	Comments and Format	Mandatory (M) or Optional (O)
<110>	Applicant	Preferably max. of 10 names; one name per line; preferable format: Surname, Other Names and/or Initials	M
<120>	Title of Invention		M
<130>	File Reference	Personal file reference	M, when filed prior to assignment of appl. number
<140>	Current Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if available
<141>	Current Filing Date	Specify as: yyyy-mm-dd	M, if available
<150>	Prior Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if applicable include priority documents under 35 USC 119 and 120
<151>	Prior Application Filing Date	Specify as: yyyy-mm-dd	M, if applicable
<160>	Number of SEQ ID NOs	Count includes total number of SEQ ID NOs	M
<170>	Software	Name of software used to create the Sequence Listing	O
<210>	SEQ ID NO:M:	Response shall be an integer representing the SEQ ID NO shown	M
<211>	Length	Respond with an integer M expressing the number of bases or amino acid residues	M

<212>	Type	Whether presented sequence molecule is DNA, RNA, or PRT (protein). If a nucleotide sequence contains both DNA and RNA fragments, the type shall be "DNA." In addition, the combined DNA/RNA molecule shall be further described in the <220> to <223> feature section.	M
<213>	Organism	Scientific name, i.e. Genus/species, Unknown or Artificial Sequence. In addition, the "Unknown" or "Artificial Sequence" organisms shall be further described in the <220> to <223> feature section.	M
<220>	Feature	Leave blank after <220>. <221-223> provide for a description of points of biological significance in the sequence.	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<221>	Name/Key	Provide appropriate identifier for feature, preferably from WIPO Standard ST.25 (1998), Appendix 2, Tables 5 and 6	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence
<222>	Location	Specify location within sequence; where appropriate state number of first and last bases/amino acids	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified

		in feature	base was used in a sequence
<223>	Other Information	Other relevant information; four lines maximum	H, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<300>	Publication Information	Leave blank after <300>	0
<301>	Authors	Preferably max of ten named authors of publication; specify one name per line; preferable format: Surname, Other Names and/or Initials	0
<302>	Title		0
<303>	Journal		0
<304>	Volume		0
<305>	Issue		0
<306>	Pages		0
<307>	Date	Journal date on which data published; specify as yyyy-mm-dd, MM-yyyy or Season-yyyy	0
<308>	Database Accession Number	Accession number assigned by database including database name	0
<309>	Database Entry Date	Date of entry in database; specify as yyyy-mm-dd or MM-yyyy	0
<310>	Patent Document Number	Document number; for patent-type citations only. Specify as, for example, US 07/999,999	0

<311>	Patent Filing Date	Document filing date, for patent-type citations only; specify as yyyy-mm-dd	0
<312>	Publication Date	Document publication date, for patent-type citations only; specify as yyyy-mm-dd	0
<313>	Relevant Residues	FROM (position) TO (position)	0
<400>	Sequence	SEQ ID NO should follow the numeric identifier and should appear on the line preceding the actual sequence	M

5. Section 1.024 is revised to read as follows:

1.024 Form and format for nucleotide and/or amino acid sequence submissions in computer readable form.

(a) The computer readable form required by 1.021(c) shall meet the following specifications:

(1) The computer readable form shall contain a single "Sequence Listing" as either a diskette, series of diskettes, or other permissible media outlined in paragraph (c) of this section.

(2) The "Sequence Listing" in paragraph (a) (1) of this section shall be submitted in American Standard Code for Information Interchange (ASCII) text. No other formats shall be allowed.

(3) The computer readable form may be created by any means, such as word processors, nucleotide/amino acid sequence editors or other custom computer programs; however, it shall conform to all specifications detailed in this section.

(4) File compression is acceptable when using diskette media, so long as the compressed file is in a self-extracting format that will decompress on one of the systems described in paragraph (b) of this section.

(5) Page numbering shall not appear within the computer readable form version of the "Sequence Listing" file.

(6) All computer-readable forms shall have a label permanently affixed thereto on which has been hand-printed or typed: the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form, the operating system used, a reference number, and an application serial number and filing date, if known.

(b) Computer readable form submissions must meet these format requirements:

(1) Computer: IBM PC/XT/AT, or compatibles, or Apple Macintosh;

(2) Operating System: MS-DOS, Unix or Macintosh;

1/29/99 1:51 PM